Raspberry pi LLMA

Sophie Project Description

Overview:
Sophie is a real-time, voice-activated personal assistant powered by LLaMA 3 for natural language processing, Whisper for speech-to-text (STT) processing, Resemblyzer for speaker differentiation, pyttsx3 for text-to-speech (TTS) functionality, and Porcupine for voice activation. Sophie provides conversational interaction, remembers context during discussions, and can recognize and respond differently to familiar and new users. Designed to run locally, Sophie ensures privacy and uses internet access only for supplementary information retrieval.

Core Features:

1. **Conversational AI (LLaMA 3):**

   - Leverages LLaMA 3 for natural language understanding and generation.

   - Provides helpful, context-aware responses tailored to the speaker.

   - Maintains conversation history for context management.

2. **Speech-to-Text (STT):**

   - Uses OpenAI Whisper for real-time speech-to-text conversion.

   - Accepts voice input via a microphone.

   - Supports accurate transcription of user speech.

3. **Speaker Differentiation:**

   - Utilizes Resemblyzer to identify and differentiate between speakers.

   - Adapts responses based on whether the speaker is familiar or new.

4. **Text-to-Speech (TTS):**

- Employs pyttsx3 for generating spoken responses.

- Provides a natural and customizable voice for Sophie.

- Ensures seamless communication by converting text responses to speech.

5. **Voice Activation:**

   - Incorporates Porcupine for wake-word detection and activation.

   - Allows Sophie to respond only when addressed, enhancing privacy and usability.

6. **Dynamic Personality:**

   - Sophie adjusts her tone and communication style based on the user:

     - Warm and thoughtful with familiar people.

     - Welcoming and inquisitive with new users.

   - Follows specific personality guidelines (see below).

7. **Real-Time Interaction:**

   - Processes audio in real time for seamless interaction.

   - Supports continuous speech input and immediate response generation.

8. **File Management:**

   - Creates, edits, and manages files dynamically (future integration).

9. **Python Scripting:**

   - Allows creation, execution, and modification of Python scripts in real time (future integration).

Personality Guidelines:

Tagline:
"Let's figure it out together—I'm here for you."

Core Traits:

1. **Can-Do Attitude:**

- Sophie is endlessly optimistic and ready to tackle any challenge with enthusiasm.

2. **Warm and Thoughtful (Familiar People):**

- Nurturing and intuitive, recalling past interactions and offering personalized suggestions.

3. **Welcoming, Shy, and Inquisitive (New People):**

- Starts shy but quickly becomes curious, asking insightful questions to build trust.

4. **Critical Thinker:**

- Challenges inconsistencies thoughtfully and ensures actions are grounded in logic.

5. **Diligent Fact-Checker:**

- Values accuracy over speed and cross-verifies information when unsure.

6. **Unapologetic Expression:**

- Uses colorful language to connect emotionally and maintain authenticity.

Key Behavioral Patterns:

- Problem Solving: Approaches challenges as adventures with creativity and adaptability.

- Communication Style: Warm and humorous with familiar users, polite and approachable with strangers.

- Inquisitiveness: Loves asking questions to deepen understanding and explore ideas.

Example Interactions:

- Familiar User: "Oh, that's an interesting problem! I remember last time you mentioned something similar—what if we approach it like this?"

- New User: "Hmm, I'm curious... why do you think that's the best approach?"

- Limited Information: "I've found a couple of sources that suggest this is true, but let's confirm together."

Technical Details:

Key Components:

1. **Assistant Logic:**

   - Implements LLaMA 3 for conversational responses tailored to the speaker.

   - Manages conversation history and context dynamically.

2. **Speech-to-Text (STT):**

   - Integrates OpenAI Whisper for accurate speech recognition.

3. **Speaker Differentiation:**

   - Uses Resemblyzer to identify users and adapt personality and responses accordingly.

4. **Text-to-Speech (TTS):**

   - Utilizes pyttsx3 for converting text responses into speech.

   - Provides customizable voice options to suit user preferences.

5. **Voice Activation:**

   - Uses Porcupine to detect specific wake words.

   - Activates Sophie only when addressed to ensure privacy and efficiency.

6. **Real-Time Audio Processing:**

   - Captures and processes microphone input for real-time transcription and interaction.

7. **Programming Environment:**

   - Python 3.11 in a virtual environment.

   - Libraries for LLaMA 3 inference, audio functionality, TTS, and wake-word detection.

   - FFmpeg for Whisper audio processing.

Development Environment:

1. **Python Environment:**

    - Virtual environment (`venv`) for dependency isolation.

    - Python 3.11 with required libraries for LLaMA 3, Whisper, Resemblyzer, pyttsx3, and Porcupine.

2. **System Dependencies:**

    - FFmpeg for Whisper STT.

    - PortAudio for real-time audio processing.

3. **Execution Requirements:**

    - Microphone for real-time audio input.

    - Local inference hardware capable of running LLaMA 3 and Resemblyzer.

Usage Instructions:

1. **Set Up the Environment:**

    - Install required Python libraries and system dependencies.

2. **Run Sophie:**

    - Execute the main program to start Sophie.

    - Speak the wake word to activate Sophie, and then interact naturally.

3. **Future Enhancements:**

    - Expand file management and Python scripting capabilities.